

(19) World Intellectual Property  
Organization  
International Bureau



(43) International Publication Date  
23 September 2004 (23.09.2004)

PCT

(10) International Publication Number  
**WO 2004/082148 A1**

(51) International Patent Classification<sup>7</sup>: **H03M 13/03**,  
13/35

[FR/FR]; c/o Société Civile SPID, 156, Boulevard Haussmann, F-75008 Paris (FR).

(21) International Application Number:  
PCT/IB2004/000682

(74) Agent: **LANDOUSY, Christian**; Société Civile SPID,  
156, Boulevard Haussmann, F-75008 Paris (FR).

(22) International Filing Date: 4 March 2004 (04.03.2004)

(81) Designated States (unless otherwise indicated, for every kind of national protection available): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, YU, ZA, ZM, ZW.

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:  
03290604.2 11 March 2003 (11.03.2003) EP

(71) Applicant (for all designated States except US): **KONINKLIJKE PHILIPS ELECTRONICS N.V.** [NL/NL];  
Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).

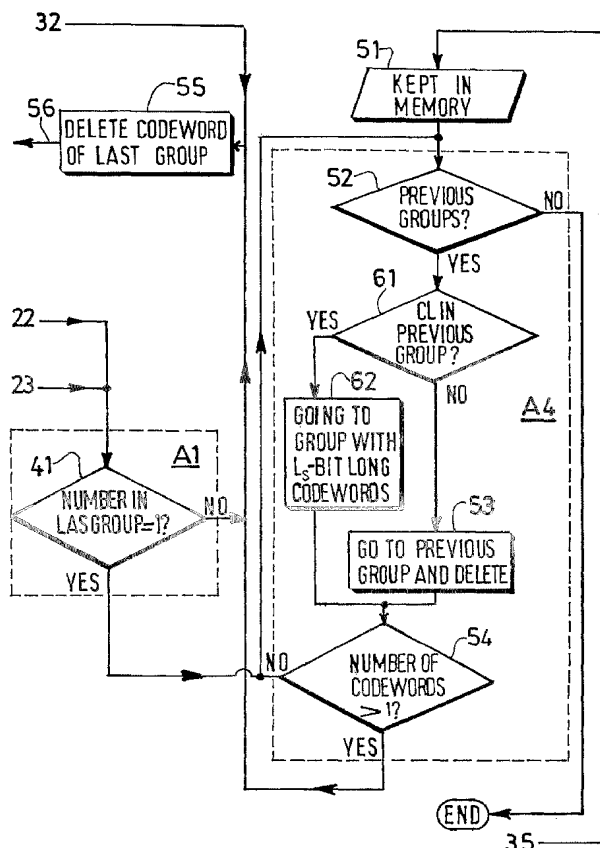
(72) Inventor; and

(75) Inventor/Applicant (for US only): **LAMY, Catherine**

(84) Designated States (unless otherwise indicated, for every kind of regional protection available): ARIPO (BW, GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW),

[Continued on next page]

(54) Title: METHOD AND DEVICE FOR BUILDING A VARIABLE-LENGTH ERROR CODE



(57) Abstract: The invention relates to a variable-length error-correcting (VLEC) code technique, in which the main steps are : defining all the needed parameters, generating a code having a fixed length L1, storing in a set W thus obtained all the possible L1-tuples distant of the minimum diverging distance d[min] from the codewords (one extra-bit being affixed at the end of all words if the new set W thus obtained is not empty), deleting all words of W that do not satisfy a distance criterion with all codewords, and verifying that all words of the final set W satisfy another distance criterion. According to the invention, it is proposed to realize the codeword deletion not anymore only in the last obtained groupe of the code, but in the group of a given length value Ls to which the algorithm will skip back to in the codeword deletion operation, which allows to go back very quickly to smaller lengths and skip many steps of the previous methods.



Eurasian (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

**Declaration under Rule 4.17:**

- as to applicant's entitlement to apply for and be granted a patent (Rule 4.17(ii)) for the following designations AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BW, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NA, NI, NO, NZ, OM, PG, PH, PL, PT, RO, RU, SC, SD, SE, SG, SK, SL, SY, TJ, TM, TN, TR, TT, TZ, UA, UG, UZ, VC, VN, YU, ZA, ZM, ZW, ARIPO patent (BW, GH, GM, KE, LS, MW, MZ, SD,

SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HU, IE, IT, LU, MC, NL, PL, PT, RO, SE, SI, SK, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG)

**Published:**

- with international search report  
— before the expiration of the time limit for amending the claims and to be republished in the event of receipt of amendments

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

# "METHOD AND DEVICE FOR BUILDING A VARIABLE-LENGTH ERROR CODE"

## FIELD OF THE INVENTION

The present invention relates to a method of building a variable length error code, said method comprising the steps of :

(1) initializing the needed parameters : minimum and maximum length of codewords  $L_1$  and  $L_{\max}$  respectively, free distance  $d_{\text{free}}$  between each codeword (said distance  $d_{\text{free}}$  being for a VLEC code C the minimum Hamming distance in the set of all arbitrary extended codes), required number of codewords S ;

(2) generating a fixed length code C of length  $L_1$  and minimal distance  $b_{\min}$ , with  $b_{\min} = \min \{b_k ; k = 1, 2, \dots, R\}$ ,  $b_k$  = the distance associated to the codeword length  $L_k$  of code C and defined as the minimum Hamming distance between all codewords of C with length  $L_k$ , and  $R$  = the number of different codeword lengths in C, said generating step creating a set W of n-bit long words distant of d ;

(3) listing and storing in the set W all the possible  $L_1$  – tuples at the distance of  $d_{\min}$  from the codewords of C (said distance  $d_{\min}$  for a VLEC code C being the minimum value of all the diverging distances between all possible couples of different-length codewords of C), and, if said set W is not empty, doubling the number of words in W by affixing at the end of all words one extra bit, said storing step therefore replacing the set W by a new one having twice more words than the previous one and the length of each one of these words being  $L_1 + 1$  ;

(4) deleting all the words of the set W that do not satisfy the  $c_{\min}$  distance with all codewords of C, said distance  $c_{\min}$  being the minimum converging distance of the code C ;

(5) in the case where no word is found or the maximum number of bits is reached, reducing the constraint of distance for finding more words ;

(6) controlling that all words of the set W are distant of  $b_{\min}$ , the found words being then added to the code C ;

(7) if the required number of codewords has not been reached, repeating the steps (1) to (6) until the method finds either no further possibility to continue or the required number of codewords ;

(8) if the number of codewords of C is greater than S, calculating on the basis of the structure of the VLEC code, the average length AL obtained by weighting

each codeword length with the probability of the source, said AL becoming the  $AL_{\min}$ , if it is lower than  $AL_{\min}$ , with  $AL_{\min}$  = the minimum value of AL, and the corresponding code structure being kept in memory.

The invention also relates to a corresponding device, allowing to carry out said building method.

## BACKGROUND OF THE INVENTION

A classical communication chain, illustrated in Fig.1, comprises, for coding the signals coming from a source S, a source coder 1 (SCOD) followed by a channel coder 2 (CCOD) and, after the transmission of the coded signals thus obtained through a channel 3, a channel decoder 4 (CDEC) and a source decoder 5 (SDEC). The decoded signals are intended to be sent towards a receiver. Variable-length codes (VLC) are classically used in source coding for their compression capabilities, and the associated channel coding techniques combat the effects of the real transmission channel (such as fading, noise, etc.). However, since source coding is intended to remove redundancy and channel coding to re-introduce it, it has been investigated how to efficiently coordinate these techniques in order to improve the overall system while keeping the complexity at an acceptable level.

Among the solutions proposed in such an approach, the variable-length error correcting (VLEC) codes present the advantage to be variable-length while providing error correction capabilities, but building these codes is rather time consuming for short alphabets (and become even prohibitive for higher length alphabets sources), and the construction complexity is also a drawback, as it will be seen.

First, some definitions and properties of the classical VLC must be recalled. A code C is a set of S codewords  $\{c_1, c_2, c_3, \dots, c_i, \dots, c_S\}$ , for each of which a length  $\ell_i = |c_i|$  is defined, with  $\ell_1 \leq \ell_2 \leq \ell_3 \leq \dots \leq \ell_i \leq \dots \leq \ell_S$  without any loss of generality. The number of different codeword lengths in the code C is called R, with obviously  $R \leq S$ , and these lengths are denoted as  $L_1, L_2, L_3, \dots, L_i, \dots, L_R$ , with  $L_1 < L_2 < L_3 < \dots < L_R$ . A variable-length code, or VLC, is then the structure denoted by  $(s_1 @ L_1, s_2 @ L_2, s_3 @ L_3, \dots, s_R @ L_R)$ , which corresponds to  $s_1$  codewords of length  $L_1$ ,  $s_2$  codewords of length  $L_2$ ,  $s_3$  codewords of length  $L_3$ , ..., and  $s_R$  codewords of length  $L_R$ . When using a VLC, the compression efficiency, for a given source, is related to the number of bits necessary to transmit symbols from said

source. The measure used to estimate this efficiency is often the average length AL of the code, i.e. the average number of bits needed to transmit a word, and said average length is given, when each symbol  $a_i$  is mapped to the codeword  $c_i$ , by the following relation (1) :

$$5 \quad AL = \sum_{i=1}^{i=S} \ell_i \cdot P(a_i) \quad (1)$$

which is equivalent to the relation (2) :

$$AL = \sum_{i=1}^R L_i \cdot \left( \sum_{j=r(i)+1}^{j=r(i+1)} P(a_j) \right) \quad (2)$$

10 where, for a data source A, the S source symbols are denoted by  $\{a_1, a_2, a_3, \dots, a_s\}$  and  $P(a_i)$  is the respective probability of occurrence of each of these symbols, with  $\sum P(a_i) = 1$  (from  $i = 1$  to  $i = S$ ). If  $AL_{\min}$  denotes the minimal value for the average length AL, it is easy to see that when  $AL_{\min}$  is reached, the symbols are indexed in such a way that  $P(a_1) \geq P(a_2) \geq P(a_3) \geq \dots \geq P(a_i) \geq \dots P(a_s)$ . In order to encode the data in such a way that the receiver can decode the coded information, the VLC  
15 must satisfy the following properties : to be non-singular (all the codewords are distinct, i.e. no more than one source symbol is allocated to one codeword) and to be uniquely decodable (i.e. it is possible to map any string of codewords unambiguously back to the correct source symbols, without any error).

20 An introduction and a presentation of different distances that are useful when reviewing some general properties of the VLC codes will then help to recall the notion of error-correcting property used in the VLEC code theory :

(a) Hamming weight and distance : if  $w$  is a word of length  $n$  with  $w = (w_1, w_2, \dots, w_n)$ , the Hamming weight of  $w$ , or simply weight, is the number  $W(w)$  of non-zero symbols in  $w$  :

$$25 \quad W(w) = \sum_{i=1}^{i=n} \frac{w_i}{\|w_i\|} \quad (3)$$

and, if  $w_1$  and  $w_2$  are two words of equal length  $n$  with  $w_i = (w_{i1}, w_{i2}, w_{i3}, \dots, w_{in})$  and  $i = 1$  or  $2$ , the Hamming distance (or, simply, distance) between  $w_1$  and  $w_2$  is the number of positions in which  $w_1$  and  $w_2$  differ (for example, for the binary case, it is easy to see that :

$$30 \quad H(w_1, w_2) = W(w_1 + w_2) \quad (4)$$

where the addition is modulo-2). However, the Hamming distance is by definition restricted to fixed-length codes, and other definitions will be defined before considering VLEC codes.

(b) let  $f_i = w_1^i w_2^i \dots w_n^i$  be a concatenation of  $n$  words of a VLEC code  $C$ , then the set  $F_N = \{f_i : |f_i| = N\}$  is called the extended code of  $C$  of order  $N$ .

(c) minimum block distance and overall minimum block distance : the minimum block distance  $b_k$  associated to the codeword length  $L_k$  of a VLEC code  $C$  is defined as the minimum Hamming distance between all distinct codewords of  $C$  with the same length  $L_k$  :

$$b_k = \min \{H(c_i, c_j) : c_i, c_j \in C, i \neq j, |c_i| = |c_j| = L_k\} \text{ for } k = 1, \dots, R \quad (5)$$

and the overall minimum block distance  $b_{\min}$  of said VLEC code  $C$ , which is the minimum block distance value for every possible length  $L_k$ , is defined by :

$$b_{\min} = \min \{b_k : k = 1, \dots, R\} \quad (6)$$

(d) diverging distance and minimum diverging distance : the diverging distance between two codewords of different length  $c_i = x_{i_1} x_{i_2} \dots x_{i_{\ell_i}}$  and

$c_j = x_{j_1} x_{j_2} \dots x_{j_{\ell_j}}$  of a VLEC code  $C$ , where  $c_i, c_j \in C$ ,  $\ell_i = |c_i|$  and  $\ell_j = |c_j|$  with  $\ell_i > \ell_j$ , is defined by :

$$D(c_i, c_j) = H(x_{i_1} x_{i_2} \dots x_{i_{\ell_i}}, x_{j_1} x_{j_2} \dots x_{j_{\ell_j}}) \quad (7)$$

i.e. it is also the Hamming distance between a  $\ell_j$  - length codeword and the  $\ell_j$  - length prefix of a longer codeword, and the minimum diverging distance  $d_{\min}$  of said VLEC code  $C$  is the minimum value of all the diverging distances between all possible couples of codewords of  $C$  of unequal length :

$$d_{\min} = \min \{D(c_i, c_j) : c_i, c_j \in C, |c_i| \neq |c_j|\} \quad (8)$$

(e) converging distance and minimum converging distance : the converging distance between two codewords of different length  $c_i = x_{i_1} x_{i_2} \dots x_{i_{\ell_i}}$  and

$c_j = x_{j_1} x_{j_2} \dots x_{j_{\ell_j}}$  of a VLEC code  $C$ , where  $|c_i| = \ell_i > |c_j| = \ell_j$ , is defined by :

$$C(c_i, c_j) = H(x_{i_{\ell_i - \ell_j + 1}} x_{i_{\ell_i - \ell_j + 2}} \dots x_{i_{\ell_i}}, x_{j_1} x_{j_2} \dots x_{j_{\ell_j}}) \quad (9)$$

i.e. it is also the Hamming distance between a  $\ell_j$  - length codeword and the  $\ell_j$  - length suffix of a longer codeword, and the minimum converging distance of said VLEC code  $C$  is the minimum value of all the converging distances between all possible couples of  $C$  of unequal length :

$$c_{\min} = \min \{C(c_i, c_j) : c_i, c_j \in C, |c_i| \neq |c_j|\} \quad (10)$$

(f) free distance : the free distance  $d_{\text{free}}$  of a code is the minimum Hamming distance in the set of all arbitrary long paths that diverge from some common state  $S_i$  and converge again in another common state  $S_j$ , with  $j > i$  :

$$d_{\text{free}} = \min \{H(f_i, f_j) : f_i, f_j \in F_N, N = 1, 2, \dots, \infty\} \quad (11)$$

Following the structure model used for a VLC, it is therefore possible to describe the structure of the VLEC code  $C$  by the notation :

$$S_1 @ L_1, b_1 ; S_2 @ L_2, b_2 ; \dots ; S_R @ L_R, b_R ; d_{\min}, c_{\min} \quad (12)$$

where there are  $s_i$  codewords of length  $L_i$  with minimum block distance  $b_i$ , for all  $i = 1, 2, \dots, R$ , (it is recalled that  $R$  is the number of different codeword lengths) and minimum diverging and converging distances  $d_{\min}$  and  $c_{\min}$ . The most important parameter of a VLEC code is its free distance  $d_{\text{free}}$ , which influences greatly its performance in terms of error-correcting capabilities, and it can be shown that the free distance of a VLEC code is bounded by :

$$d_{\text{free}} \geq \min (b_{\min}, d_{\min} + c_{\min}) \quad (13)$$

These definitions being recalled, the state-of-the-art in VLEC codes construction will be now described more easily. The first types of VLEC codes, called  $\alpha$ -prompt codes and introduced in 1974, and an extension of this family, called  $\alpha_{t_1, t_2, \dots, t_R}$ -prompt codes, have both the same essential property : if one denotes by  $\alpha(c_i)$  the set of words that are closer to  $c_i$  than to any codeword  $c_j$ , with  $j \neq i$ , no sequence in  $\alpha(c_i)$  is a prefix of a sequence in another  $\alpha(c_i)$ . The construction of these codes is very simple, and the construction algorithm is adjustable by the number of codewords at each length, which makes possible to find the best prompt code for a given source and a given  $d_{\text{free}}$ . However, this best code performs poorly in terms of compression performance.

A more recent construction, allowing the construction of a VLEC code from the generator matrix of a fixed-length linear block code, was proposed in the document "Variable-length error-correcting codes" by V. Buttigieg, Ph.D. Thesis, University of Manchester, England, 1995. Called code-anticode construction, this algorithm relies on line combinations and column permutations to form an anticode at the rightmost column. Once the code-anticode generator matrix is obtained, the VLEC code is simply obtained by a matrix multiplication.

This technique has however several drawbacks. First, there is no explicit method to find the needed line combinations and column permutations to obtain the anticode. Moreover, the construction does not take into account the source statistics and, consequently, often reveals itself sub-optimal (one can find a code with smaller average length by a post-processing on the VLEC code). In the same document, the author has then proposed an improved method, called Heuristic method, that is based on a computer search for building a VLEC code giving the better known compression rate for a specified source and a given protection against errors, i.e. a code C with specified overall minimum block, diverging and converging distances (and hence a minimum value for  $d_{\text{free}}$ ) and with codeword lengths matched to the source statistics so as to obtain a minimum average codeword length for the chosen free distance and the specified source (in practice, one takes :  $b_{\text{min}} = d_{\text{min}} + c_{\text{min}} = d_{\text{free}}$ , and :  $d_{\text{min}} = \lfloor d_{\text{free}}/2 \rfloor$ ).

The main steps of this Heuristic method, which uses the following parameters : minimum length  $L_1$  of codewords, maximum length  $L_{\text{max}}$  of codewords, free distance  $d_{\text{free}}$  between each codeword, number S of codewords required, are now described with reference to the flowcharts of Figs.2 to 4.

To start the computer search ("Start"), all the needed parameters must be first specified :  $L_1$  (the minimum codeword length, which must be at least equal to or greater than the minimum diverging distance required),  $L_{\text{max}}$  (the maximum codeword length), the different distances between codewords ( $d_{\text{free}}$ ,  $b_{\text{min}}$ ,  $d_{\text{min}}$ ,  $c_{\text{min}}$ ), and S (the number of codewords required by the given source), and some relations are set when choosing these parameters :

$$\begin{aligned} L_1 &\geq d_{\text{min}} \\ b_{\text{min}} &= d_{\text{free}} \\ d_{\text{min}} + c_{\text{min}} &= d_{\text{free}} \end{aligned}$$

The first phase of the algorithm, referenced 11, is then performed : it consists in the generation of a fixed length code (put initially in C) of length  $L_1$  and minimal distance  $b_{\text{min}}$ , with a maximum number of codewords. This phase is in fact an initialization, performed for instance by means of an algorithm such as the greedy algorithm (GA), presented in Fig.5, or the majority voting algorithm (MVA), presented in Fig.7, or a new proposed variation, denoted by GAS (Greedy Algorithm by Step), which consists of a variation of the two above mentioned ones. The GAS consists in the search method used in the GA, where instead of deleting half of the



codewords, only the last codeword of the group is deleted. These two algorithms are useful to create a set  $W$  of  $n$ -bit long words distant of  $d$  (in practice, it may be noted that the MVA finds more words than the GA, but it asks too much time for only a small improvement of the compression capacity, as shown in the tables of Figs.6 and 8, which compare, respectively for the GA and for the MVA, the best code structures obtained with different values of  $d_{\text{free}}$  for the 26-symbol English source defined in the table of Fig.9.

The second phase of the algorithm, corresponding to the elements referenced 21 to 24 (21+22 = operation "A0" ; 23+24 = operation "A2") in Fig.2, consists in listing and storing (step 21) in a set called  $W$  all the possible  $L_1$  - tuples at the distance of  $d_{\text{min}}$  from the codewords in  $C$ . If  $d_{\text{min}} \geq b_{\text{min}}$ , then  $W$  is empty. If this set  $W$  of all the words satisfying the minimum diverging distance to the current code is not empty (reply NO to the test 22 :  $|W| = 0$  ?), the number of words in  $W$  is doubled by increasing the length of the words by one bit by affixing first a "0" and then a "1" to the rightmost position of all the words in  $W$  (step 24), except if the maximum number of bits is exceeded (reply YES to the test 23). At the output of said step 24, this modified set  $W$  has twice more words than the previous  $W$ , and the length of each one is  $L_1 + 1$ .

The third phase of the algorithm, corresponding to the elements 31 to 35 (= operation "A3" in Fig.2), consists in deleting (step 31) all the words of set  $W$  that do not satisfy the  $c_{\text{min}}$  distance (minimum converging distance) with all the codewords of  $C$  (i.e. in keeping and storing in a new  $W$  only the words which satisfy said minimum converging distance, the other ones being discarded). At this point, the new set  $W$  is a set of words which, when compared to the codewords of  $C$ , satisfy the required minimum diverging and converging distances (both  $d_{\text{min}}$  and  $c_{\text{min}}$  distances) with the codewords of  $C$ . If that new set  $W$  is not empty (reply NO to the test 32 :  $|W| = 0$  ?) one selects in  $W$  (step 33) the maximum number of words to satisfy the minimum block distance, in order to ensure that all the words of the set  $W$ , being of the same length, have a minimum distance at least equal to  $b_{\text{min}}$ . At the end of this step 33, realized with the GA or the MVA (note that in this case, the initial set used for the GA or the MVA is the current  $W$  and not a  $n$ -tuples set), the words thus obtained are added (step 34) to the codewords already in  $C$ .

If no word is found (i.e.  $W$  is empty) at the end of the step 21 (reply YES to the test 22 :  $|W| = 0$  ?) or if the maximum number of bits is reached or exceeded (reply YES to the test 23), one enters the fourth phase of the algorithm (steps 41 to 46, illustrated in Fig.3 and also designated by the operation "A1" in said figure), which is used in order to unjam the process by inserting more liberty of choice, more particularly by affixing to all words in  $W$  extra bits (several bits at the same time) such that the new group contains more bits than the old one. If there are enough codewords in the last group (successive tests 41 and 42, for verifying the number of codewords in the last group, and if there are previous groups), some of them are deleted from said group (as described above), such deletions allowing to reduce the distance constraint and to find more codewords than before. As a matter of fact, the classical Heuristic method thus described begins with the maximum of codewords with the short length, maps them with the high probability symbols and tries to obtain a good compression rate, but sometimes the size of the small lengths sets are incompatible with the required number of codewords  $S$ . In this optic, easing a few codewords provides more freedom degrees and allows to reach a position where the initial requirements on distance and number of symbols for the code can be met. This deletion process is repeated until it remains a maximum of one codeword for each length. If  $W$  is empty at the end of the step 31 (reply YES to the test 32 :  $|W| = 0$  ?), the steps 23, 24, 31, 32 are repeated. If the required number of codewords has not been reached (reply NO to the test 35 provided at the end of this third phase), the steps 21 to 24 and 31 to 35 must be repeated until said steps find that either there are no further possible words to be found or the required number of codewords is reached.

If said required number of codewords has been reached (i.e. the number of codewords of  $C$  is equal to or greater than  $S$  (reply YES to the test 35), the structure of the VLEC code thus obtained is used in a fifth part, including the steps 51 to 56 (illustrated in Fig.4, and also designated by the operation "A4" in said figure), in order to calculate the average length  $AL$ . This is done by weighting each codeword length with the probability of the source, and comparing it to the current best one. If said average length  $AL$  of this VLEC code is lower than the minimized value of  $AL$  ( $= AL_{\min}$ ), this  $AL$  becomes the  $AL_{\min}$ , and this new  $AL$  value and the corresponding code structure are kept in the memory (step 51). These steps 51 and following (fifth part ; operation "A4") allow to come back, within the algorithm, towards previous

groups, while the other phases of said algorithm are always performed on the current group. The stepsize for such a feedback operation is one, i.e. this feedback action can be considered as exhaustive.

To continue this search of the best VLEC code, it is necessary to avoid keeping the same structure, which would lead to a loop in the algorithm. The last added group of the current code is deleted (steps 52, 53), the deletion of shorter length codewords allowing to find more longer length codewords (test 54 : number of codewords in group greater than 1 ?), and some codewords (half the amount for the GVA ; the "best" one for the MVA) of the previous group are deleted (step 55), in order to re-loop (step 56) the algorithm at the beginning of the step 21 (see Fig.2) and find different VLEC structures (the number of deleted codewords depends on which method is used for selecting the words : if the GA method is used and one wants to obtain a linear code, it is necessary to delete half of the codewords, while with the MVA method only one codeword, the best one, is deleted, i.e. the one that allows to find the more codewords in the next group).

However, the Heuristic method thus described often considers very unlikely code structures or proceeds with such a care (in order not to miss anything) that a great complexity is observed in the implementation of said method, which moreover is rather time consuming and can thus become prohibitive. It has therefore been proposed in a European patent application filed on October 23, 2002, with the filing number 02292624.0 (PHFR020110), an improved construction method with which it is possible to gain in complexity by avoiding these drawbacks, said method of building a variable length error code comprising, more precisely, the steps of :

(1) initializing the needed parameters : minimum and maximum length of codewords  $L_1$  and  $L_{\max}$  respectively, free distance  $d_{\text{free}}$  between each codeword (said distance  $d_{\text{free}}$  being for a VLEC code  $C$  the minimum Hamming distance in the set of all arbitrary extended codes), required number of codewords  $S$  ;

(2) generating (step 11) a fixed length code  $C$  of length  $L_1$  and minimal distance  $b_{\min}$ , with  $b_{\min} = \min \{b_k ; k = 1, 2, \dots, R\}$ ,  $b_k$  = the distance associated to the codeword length  $L_k$  of code  $C$  and defined as the minimum Hamming distance between all codewords of  $C$  with length  $L_k$ , and  $R$  = the number of different codeword lengths in  $C$ , said generating step 11 creating a set  $W$  of  $n$ -bit long words distant of  $d$  ;

(3) listing and storing (step 21) in the set  $W$  all the possible  $L_1$  – tuples at the distance of  $d_{\min}$  from the codewords of  $C$  (said distance  $d_{\min}$  for a VLEC code  $C$  being the minimum value of all the diverging distances between all possible couples of different-length codewords of  $C$ ), and, if said set  $W$  is not empty, doubling the number of words in  $W$  by affixing at the end of all words one extra bit, said storing step therefore replacing the set  $W$  by a new one having twice more words than the previous one and the length of each one of these words being  $L_1 + 1$  ;

(4) deleting (step 31) all the words of the set  $W$  that do not satisfy the  $c_{\min}$  distance with all codewords of  $C$ , said distance  $c_{\min}$  being the minimum converging distance of the code  $C$  ;

(5) in the case where no word is found or the maximum number of bits is reached, reducing (step 41) the constraint of distance for finding more words ;

(6) controlling that all words of the set  $W$  are distant of  $b_{\min}$ , the found words being then added to the code  $C$  (step 34) ;

(7) if (step 35) the required number of codewords has not been reached, repeating the steps (1) to (6) (i.e. the steps 21 to 35) until the method finds either no further possibility to continue or the required number of codewords ;

(8) if the number of codewords of  $C$  is greater than  $S$ , calculating (phase A4), on the basis of the structure of the VLEC code, the average length  $AL$  obtained by weighting each codeword length with the probability of the source, said  $AL$  becoming the  $AL_{\min}$ , if it is lower than  $AL_{\min}$ , with  $AL_{\min}$  = the minimum value of  $AL$ , and the corresponding code structure being kept in memory ; said building method being moreover such that at most one bit is added at the end of each word of the set  $W$ .

Simulations show that, with the classical Heuristic method, almost none of the obtained best codes has a hole (i.e. a length jump in its structure length), while it is considered, in the cited document, that most good codes do not have jump of length and, therefore, that the set of examined VLEC codes can be reduced accordingly (which reduces the simulation time and the complexity of implementation of the method, without modifying much the  $AL$ ). Following this hypothesis, the method is, in the EP document cited above, modified by avoiding to add more than one bit at the end of each word of the set  $W$ . In this new structure, the circuit 45 has been cancelled, which means that it is no longer allowed to affix several extra bits to the words, but only one, which leads, with respect to Fig.3, to a

simpler phase A1 (Fig.11). The corresponding implementation (improved Heuristic construction method, also called "noHole optimization method") is illustrated in Figs.10 and 11, which show the two parts of a flowchart corresponding to said method (the elements that are identical to the ones observed in Figs.2 to 4 being designated with the same references). With respect to the flowchart of Figs.2 to 4, the parts that, with respect to the classical Heuristic technique, are useless for the implementation of the improved method have been cancelled :

(a) if W is empty at the end of the step 31 (reply YES to the test 32 :  $|W|=0$  ?), the next phase is now (see Fig.10) not the repetition of the steps (23, 24, 31, 32), but, according to said "noHole" method, the establishment (in place of said repetition) of a direct connection 91 towards the input of the circuit carrying out the operation 55 (deletion of some codewords, or of the best one, before a global repetition of the steps 21 to 24 and 31 to 35), said operation 55 being then, as previously, followed by the operations 21 and following.

(b) the fourth phase of the method is now reduced to one step, the operation 41, which is the test "Number of codewords in last group = 1 ?" . If the reply is NO, a direct link is established with the input of the step 55 (connection 91), in view of carrying out said operation 55, and then the operations 21 and following. If the reply is YES, a connection 92 is established with the input of the set of operations 52 to 54.

The results thus obtained are presented in the table of Fig.12 for the 26 symbol English source when using the GAS method for selecting codewords. It can be seen, when comparing with results presented in Fig.13, that although the result is not completely optimal for  $d_{\text{free}} = 3$  (the code structure has a hole at length  $L = 11$ ), the AL rise is really acceptable when one considers that there is both strictly no degradation for the other  $d_{\text{free}}$  values and a gain of time between 2,5 and 4. The same remarks can be applied when comparing the present solution with the ones obtained in Fig.7, where the MVA complexity effect is clear. Similarly, applying the noHole optimisation with the GA method for selecting codewords leads to a time gain at the only expense of a slight AL rise for  $d_{\text{free}}=3$ . Finally, Fig.5 shows on the other hand that the current solution offers better AL for an acceptable gain of time, the noHole optimisation compensating almost entirely the complexity induced by the GAS. However, with the method thus described in the EP document, there are cases where there are too many small length codewords in the generated VLEC code.

## SUMMARY OF THE INVENTION

It is therefore an object of the invention to propose an improved construction method with which it is possible to go back very quickly to smaller lengths and skip many algorithm steps (in cases where the algorithm would have painfully got stuck with too many codewords of small lengths).

To this end, the invention relates to a method such as defined in the introductory part of the description and which is moreover characterized in that the deletion is realized not only in the last obtained group but also in the group of a given length value in order to go back very quickly to smaller lengths.

According to a possible improved implementation of the invention, the invention relates to a similar method (i.e. also such as defined in the introductory part of the description), but which is now preferably characterized in that at most one bit is added at the end of each word of the set W and the deletion is moreover realized not only in the last obtained group but also in the group of a given length value in order to go back very quickly to smaller lengths.

It is also an object of the invention to propose a device for carrying out said construction method.

To this end, the invention relates to a device for carrying out a variable length error code building method according to anyones of the two solutions thus proposed.

## BRIEF DESCRIPTION OF THE DRAWINGS

The present invention will now be described, by way of example, with reference to the accompanying drawings in which :

- Fig.1 depicts a conventional communication channel ;

- Figs 2 to 4 are the three parts of a single flowchart illustrating the main steps of a conventional method used for building a VLEC code (and called Heuristic method) ;

- Fig.5 illustrates an algorithm (called greedy algorithm, or GA) used for the initialization of the method of Figs. 2 to 4, and Fig.6 is a table giving various VLEC codes for a source constructed with the Heuristic construction using said algorithm of Fig.5 ;

- Fig.7 illustrates another algorithm (called majority voting algorithm, or MVA) used for the initialization of the method of Figs. 2 to 4, and Fig.8 is another

table giving various VLEC codes for a source constructed with the Heuristic construction using said algorithm of Fig.7 ;

- Fig.9 is a table giving for the 26-symbol English source the correspondence between the source symbol and its probability ;

- Figs 10 and 11 are the two parts of a single flowchart illustrating an implementation of an improvement of the conventional method illustrated in Figs.2 to 4 ;

- Fig.12 is another table giving various VLEC codes for the same 26 symbol English source as considered in the tables of Figs.6 and 8 and using the GAS ;

- Fig.13 is another table giving various VLEC codes for the same source as in Fig.12 and using both the GAS previously mentioned and the building method according to the invention ;

- Fig.14 shows the modification of the flowchart of Fig.11 when the method according to the invention is carried out ;

- Fig.15 is a table illustrating the results obtained for the 26-symbol English source when the method according to the invention is carried out ;

- Fig.16 is a table illustrating similarly the results obtained for another source.

## DETAILED DESCRIPTION OF THE INVENTION

Considering the evolution of the codewords stack in the classical Heuristic process previously recalled, it may be proposed to perform the codeword deletion not only in the last obtained codewords group, but more generally in any given length value group, in order to make possible to go back directly to smaller lengths, i.e. to skip many algorithm steps in cases where there are too many small length codewords.

Denoting by  $L_s$  (with s for : skip) the length to which the algorithm will skip back to in the codeword deletion stage, it is therefore proposed, in the present so-called  $L_s$  optimization, to skip parts of the original Heuristic algorithm by carefully jumping to lower lengths when looking for codewords to be deleted (however, when the considered codewords group length  $L$  is smaller than  $L_s$ , the classical method is obviously applied and the deletion is done within the group of length  $L$ ). With respect to the previous fifth part (phase A4) of the "noHole" and "Heuristic" methods, the process is no longer exhaustive : it means that it is possible to come back more quickly towards previous groups, up to a target length  $L_s$  in order to

avoid to spend too much time with said feedback operations. As indicated, the classical method is however maintained for lengths lower than  $L_s$ . The lengths comprised between  $L_1$  and  $L_s$  are consequently called "free lengths", i.e. lengths with a freedom degree, as they are decremented one by one in the search process (when the number of free lengths grows up, the simulation time also increased, exponentially).

The method according to the invention is called " $L_s$  optimization" and is depicted in the flowchart formed by the association of Fig.10 (unchanged part of the previous method, the so-called noHole optimization method) and Fig.14 (modified part of the noHole optimization method), said Fig.14 being adapted from Fig.11 by means of an introduction of the technical measures proposed above.

According to said measures, the last added group of the current code is deleted, but only if (test 61) the codeword length CL of this previous group is lower (reply NO to the test 61) than or equal to  $L_s$  (the steps that follow the test 61 are then the same as previously : steps 53, 54, and 55 or 52 at the output of the test 54). If said codeword length is greater than  $L_s$  (reply YES to the test 61), an additional step 62 is provided for going to group with  $L_s$ -bit long codewords and deleting all groups with more than  $L_s$  bits. At the output of the step 62, the same steps 54, 55 as previously are provided.

In practice, simulation results show that very good compression rates can be obtained for  $L_s < L(\max)$ , where  $L(\max)$  is the maximal authorized codeword length (it can be noted that increasing the value of  $L_s$  results in an improvement of the AL value until a constant floor – the best value – is reached, and this behaviour then suggests a possible dynamic choice of  $L_s$ , starting with  $L_s = L_1$  and incrementing it until said floor is reached). The results obtained for the 26-symbol English source are presented in the table of Fig.15 (crosses indicate impossible values for  $L_s$ , which must belong to the interval  $[L_1, L(\max)]$ , and the mention "no code" corresponds to the cases where a too short value of  $L_s$  resulted in a failure of the algorithm, i.e. no VLEC code verifying the initial parameters could be found). Comparing those results with the ones of the table of Fig.12 leads to observe a noticeable time gain, from about 5 at  $d_{\text{free}} = 3$  to about 2500 at  $d_{\text{free}} = 7$ . The gain in computation time offered by the present  $L_s$  optimization algorithm also allows to find VLEC codes for the 208-symbol MPEG-4 source, as illustrated in the table of Fig.16, where the



results have been obtained using an algorithm including both the GAS and the noHole and  $L_s$  optimizations.

It must be understood that the method according to the invention is not limited to the aforementioned implementation. Another implementation can for instance correspond to the addition of the features of the "noHole" method and the present "L<sub>s</sub> optimization" method : in this case, the method according to the invention is then such that at most one bit is added at the end of each word of the set W and, in the same time, the deletion is realized not only in the last obtained group but also in the group of a given length value in order to go back very quickly to smaller lengths.

5

10

CLAIMS :

1. A method of building a variable length error code, said method comprising the steps of :

(1) initializing (phase 0) the needed parameters : minimum and maximum length of codewords  $L_1$  and  $L_{\max}$  respectively, free distance  $d_{\text{free}}$  between each codeword (said distance  $d_{\text{free}}$  being for a VLEC code C the minimum Hamming distance in the set of all arbitrary extended codes), required number of codewords S ;

(2) generating (phase 1) a fixed length code C of length  $L_1$  and minimal distance  $b_{\min}$ , with  $b_{\min} = \min \{b_k ; k = 1, 2, \dots, R\}$ ,  $b_k$  = the distance associated to the codeword length  $L_k$  of code C and defined as the minimum Hamming distance between all codewords of C with length  $L_k$ , and  $R$  = the number of different codeword lengths in C, said generating step creating a set W of n-bit long words distant of d ;

(3) storing (phase 2) in the set W all the possible  $L_1$  - tuples distant of  $d_{\min}$  from the codewords of C (said distance  $d_{\min}$  for a VLEC code C being the minimum value of all the diverging distances between all possible couples of different-length codewords of C), and, if said set W is not empty, affixing at the end of all words one extra bit, said storing step replacing the set W by a new one having twice more words than the previous one and the length of each one of these words being  $L_1 + 1$  ;

(4) deleting (phase 3) all the words of the set W that do not satisfy the  $c_{\min}$  distance with all codewords of C, said distance  $c_{\min}$  being the minimum converging distance of the code C ;

(5) in the case where no word is found or the maximum number of bits is reached, reducing (phase A1) the constraint of distance for finding more words ;

(6) controlling that all words of the set W are distant of  $b_{\min}$ , the found words being then added to the code C ;

(7) if the required number of codewords has not been reached, repeating the steps (1) to (6) until the method finds either no further possibility to continue or the required number of codewords ;

(8) if the number of codewords of C is greater than S, calculating, on the basis of the structure of the VLEC code, the average length AL obtained by weighting each codeword length with the probability of the source, said AL becoming the  $AL_{\min}$  if it is lower than  $AL_{\min}$ , with  $AL_{\min}$  = the minimum value of AL, and the corresponding code structure being kept in memory ;

said building method being such that the deletion is realized not only in the last obtained group but also in the group of a given length value in order to go back very quickly to smaller lengths.

2. A method of building a variable length error code, said method comprising the steps of :

(1) initializing (phase 0) the needed parameters : minimum and maximum length of codewords  $L_1$  and  $L_{\max}$  respectively, free distance  $d_{\text{free}}$  between each codeword (said distance  $d_{\text{free}}$  being for a VLEC code C the minimum Hamming distance in the set of all arbitrary extended codes), required number of codewords S ;

(2) generating (phase 1) a fixed length code C of length  $L_1$  and minimal distance  $b_{\min}$ , with  $b_{\min} = \min \{b_k ; k = 1, 2, \dots, R\}$ ,  $b_k$  = the distance associated to the codeword length  $L_k$  of code C and defined as the minimum Hamming distance between all codewords of C with length  $L_k$ , and  $R$  = the number of different codeword lengths in C, said generating step creating a set W of n-bit long words distant of d ;

(3) storing (phase 2) in the set W all the possible  $L_1$  - tuples distant of  $d_{\min}$  from the codewords of C (said distance  $d_{\min}$  for a VLEC code C being the minimum value of all the diverging distances between all possible couples of different-length codewords of C), and, if said set W is not empty, affixing at the end of all words one extra bit, said storing step replacing the set W by a new one having twice more words than the previous one and the length of each one of these words being  $L_1 + 1$  ;

(4) deleting (phase 3) all the words of the set W that do not satisfy the  $c_{\min}$  distance with all codewords of C, said distance  $c_{\min}$  being the minimum converging distance of the code C ;

(5) in the case where no word is found or the maximum number of bits is reached, reducing (phase A1) the constraint of distance for finding more words ;

(6) controlling that all words of the set W are distant of  $b_{\min}$ , the found words being then added to the code C ;

(7) if the required number of codewords has not been reached, repeating the steps (1) to (6) until the method finds either no further possibility to continue or the required number of codewords ;

(8) if the number of codewords of C is greater than S, calculating, on the basis of the structure of the VLEC code, the average length AL obtained by weighting each codeword length with the probability of the source, said AL becoming the

$AL_{\min}$  if it is lower than  $AL_{\min}$ , with  $AL_{\min}$  = the minimum value of  $AL$ , and the corresponding code structure being kept in memory ;

said building method being such that at most one bit is added at the end of each

word of the set  $W$ , and the deletion is moreover realized not only in the last obtained group but also in the group of a given length value in order to go back very quickly to smaller lengths.

3. A device for carrying out a variable length error code building method according to anyones of claims 1 and 2.

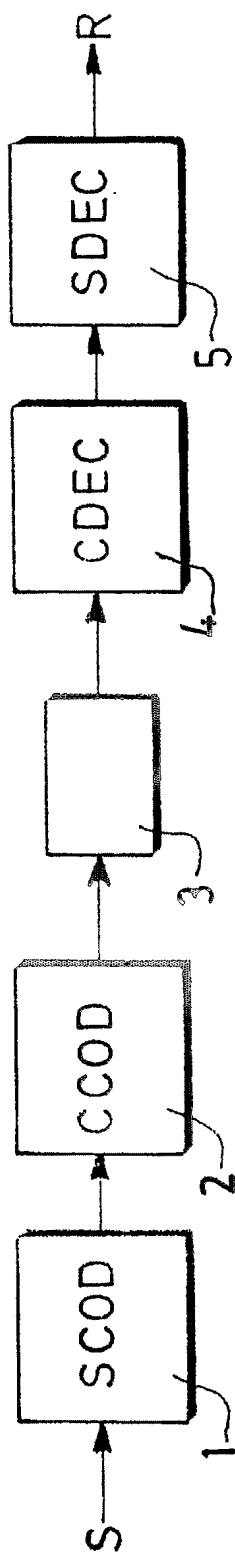
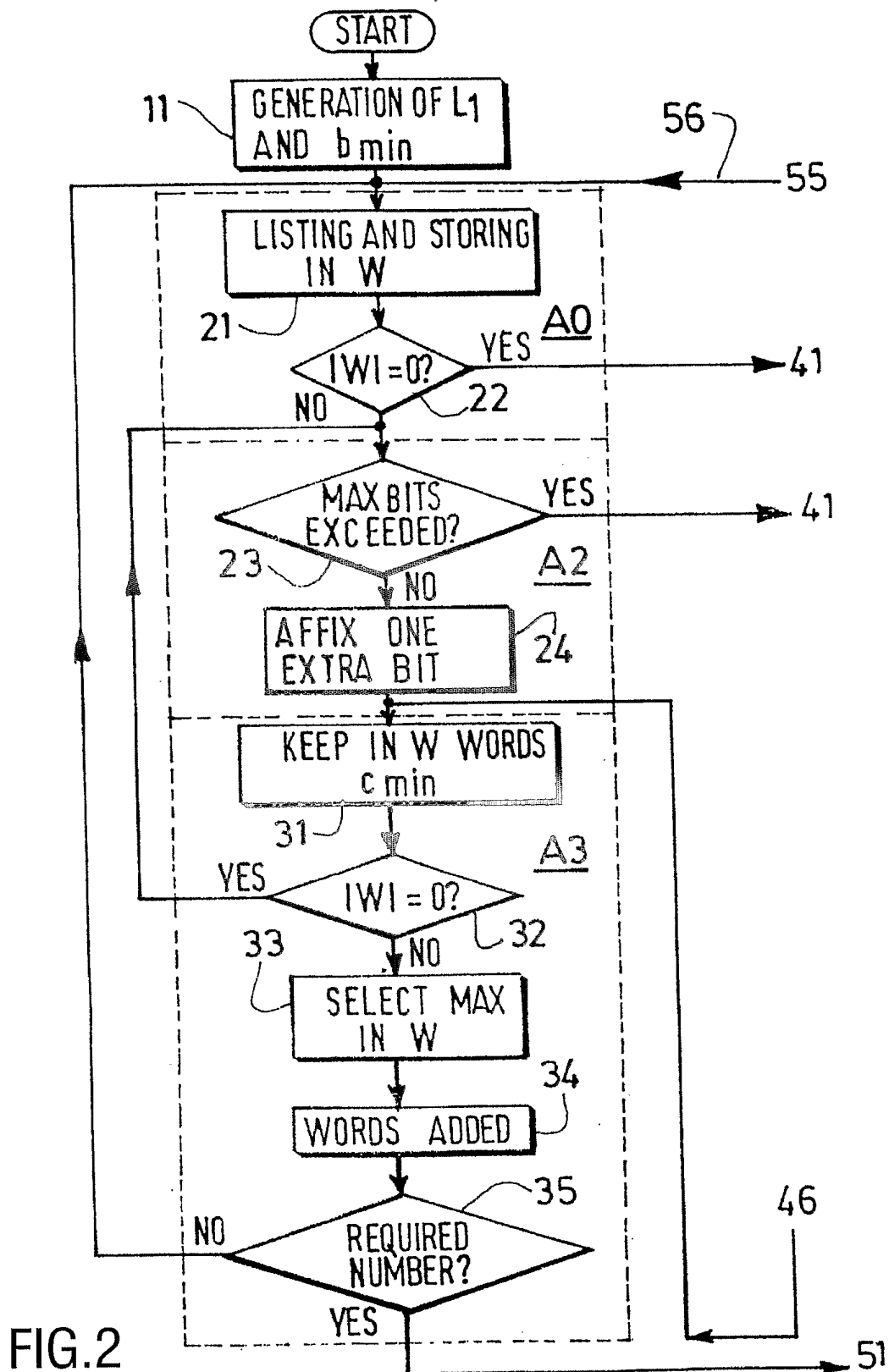


FIG. 1



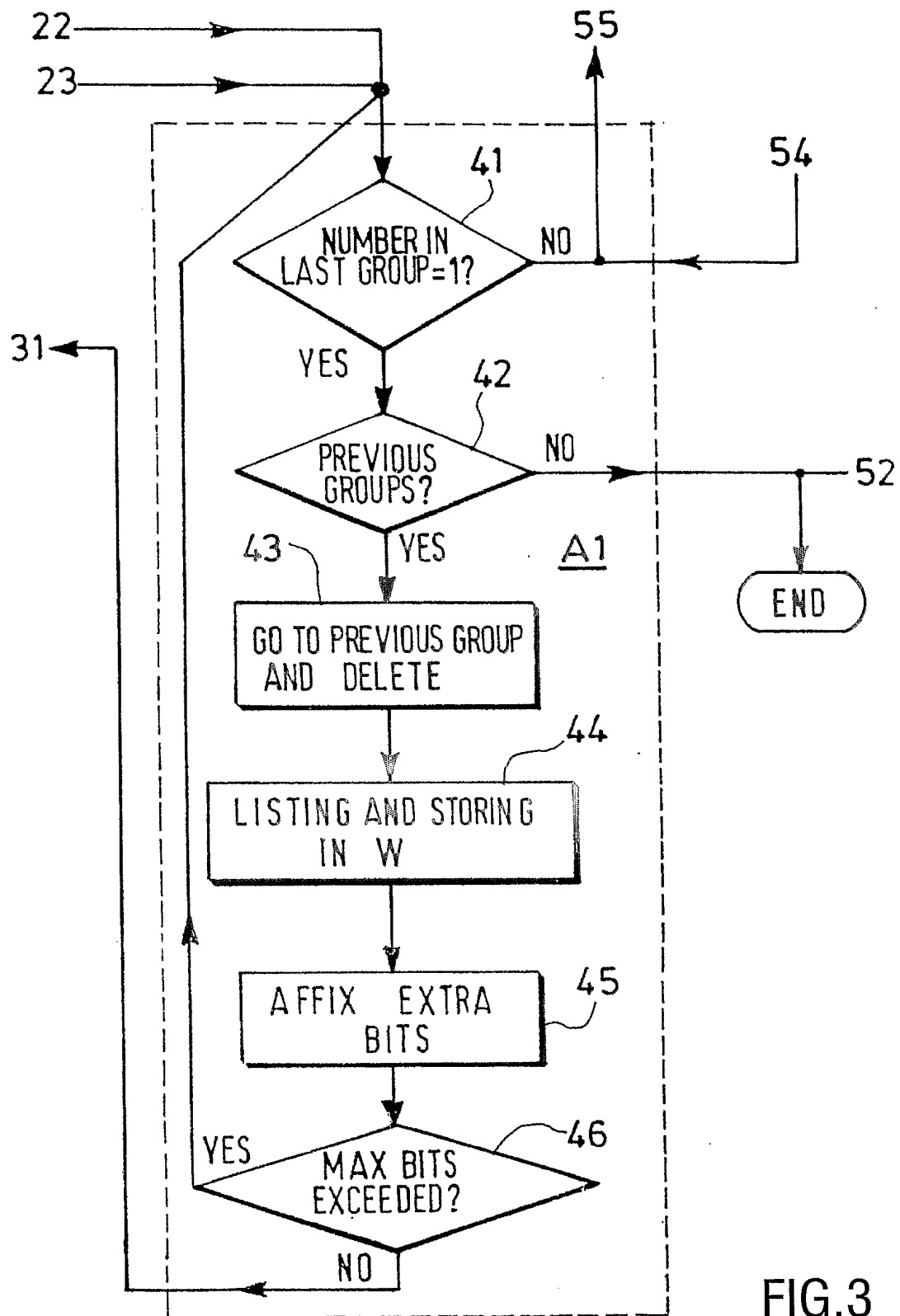


FIG.3

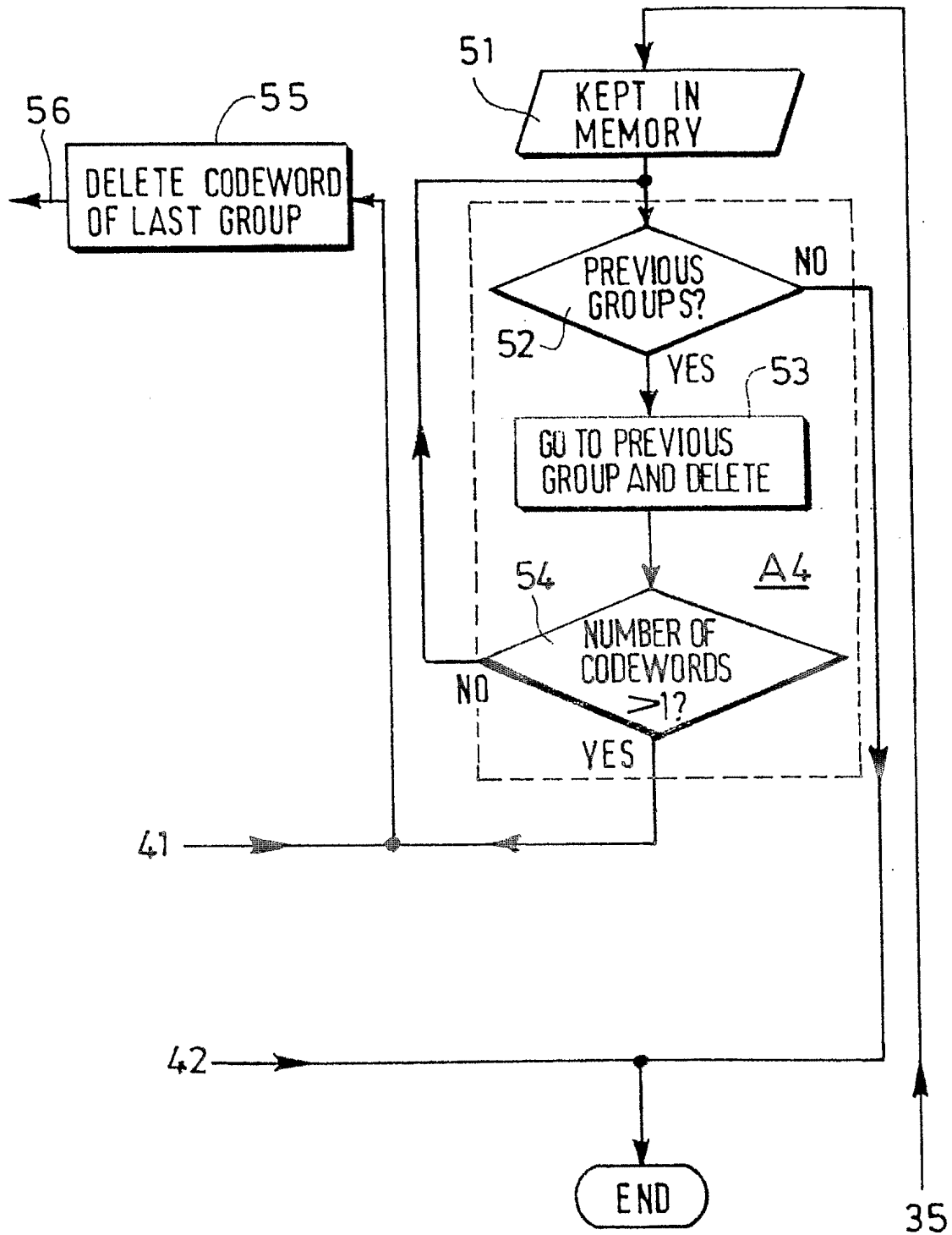


FIG. 4



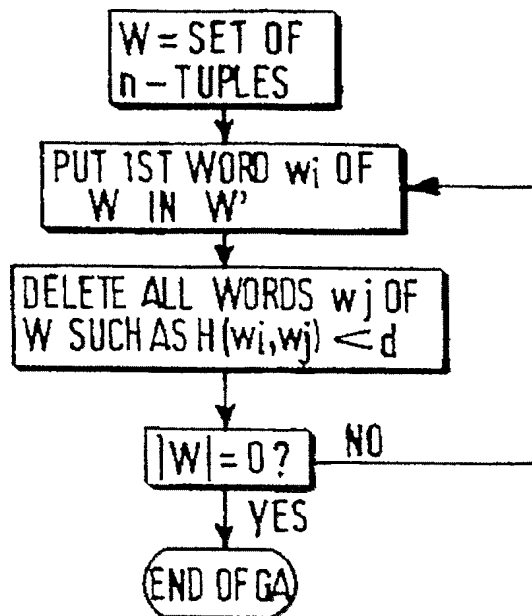


FIG.5

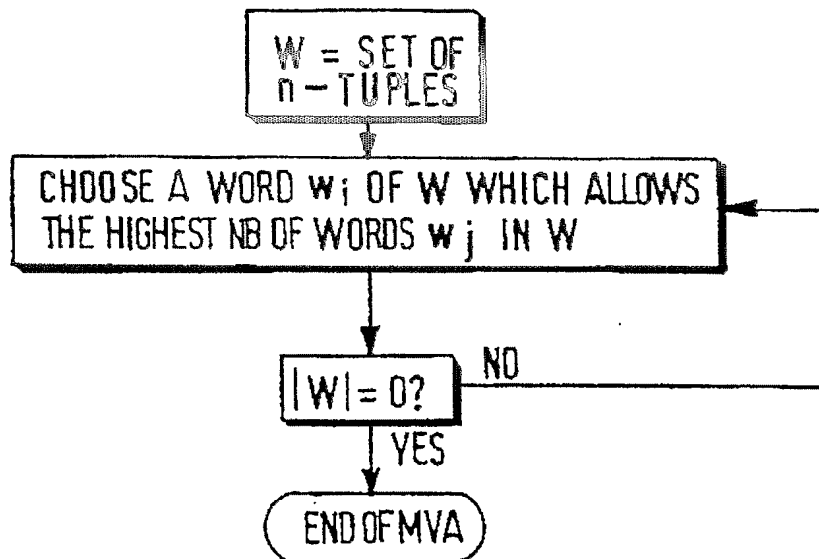


FIG.7

dfree	AL	TIME	CODE STRUCTURE
3	6,4946	4s	(1@ 4,-;1@ 5,-;3@ 6,3;8@ 7,3;4@ 8,3;3@ 9,3;1@ 10,-;2@ 11,4;1@ 12,-;1@ 13,-;1@ 14,-;2,1)
5	8,5061	15s	(1@ 6,-;2@ 7,5;3@ 8,5;4@ 9,5;4@ 10,5;5@ 11,5;5@ 12,5;2@ 13,5;3,2)
7	10,79	2min	(1@ 8,-;2@ 9,7;2@ 10,7;3@ 11,7;4@ 12,7;4@ 13,7;6@ 14,7;4@ 15,7;4,3)

FIG.6

d <sub>free</sub>	A L	TIME	CODE STRUCTURE
3	63038	50 s	(1@4,-;2@5,4;4@6,3;4@7,3;6@8,3;5@9,3;2@10,4;2@11,3;2,1)
5	84752	16 min	(1@6,-;1@7,-;4@8,5;5@9,5;5@10,5;5@11,5;5@12,5;3,2)
7	107385	14 h	(1@7,-;1@8,-;1@9,-;1@10,-;3@11,7;4@12,7;5@13,7;5@14,7;5@15,7;4,3)

FIG.8

SOURCE SYMBOL	PROBABILITY
<i>e</i>	0,1270
<i>t</i>	0,0906
<i>a</i>	0,0817
<i>o</i>	0,0751
<i>i</i>	0,0697
<i>n</i>	0,0674
<i>s</i>	0,0633
<i>h</i>	0,0609
<i>r</i>	0,0599
<i>d</i>	0,0425
<i>l</i>	0,0403
<i>c</i>	0,0278
<i>u</i>	0,0276
<i>m</i>	0,0241
<i>w</i>	0,0236
<i>f</i>	0,0223
<i>g</i>	0,0202
<i>y</i>	0,0197
<i>p</i>	0,0193
<i>b</i>	0,0149
<i>v</i>	0,0098
<i>k</i>	0,0077
<i>j</i>	0,0015
<i>x</i>	0,0015
<i>q</i>	0,0010
<i>z</i>	0,0007

FIG.9

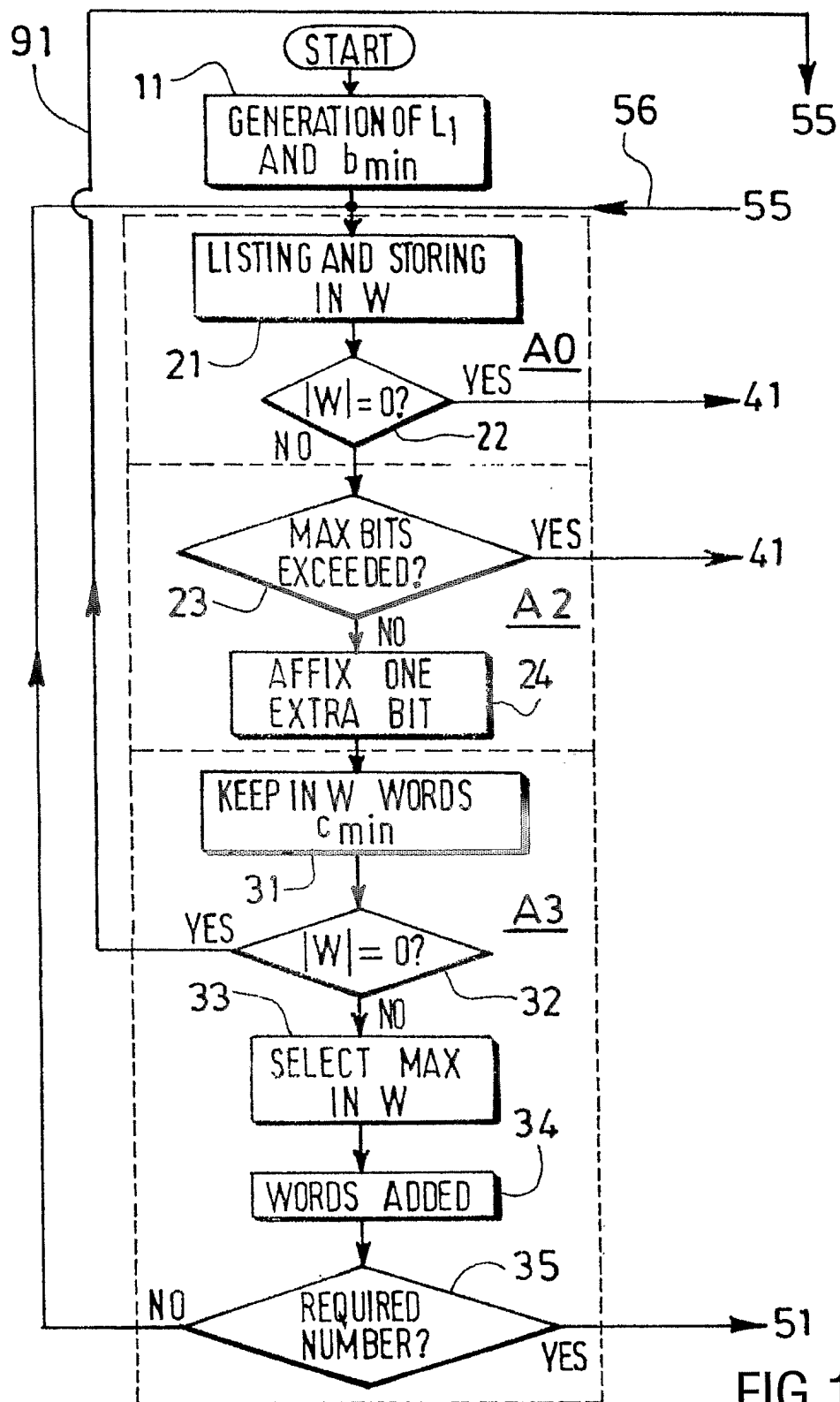


FIG.10

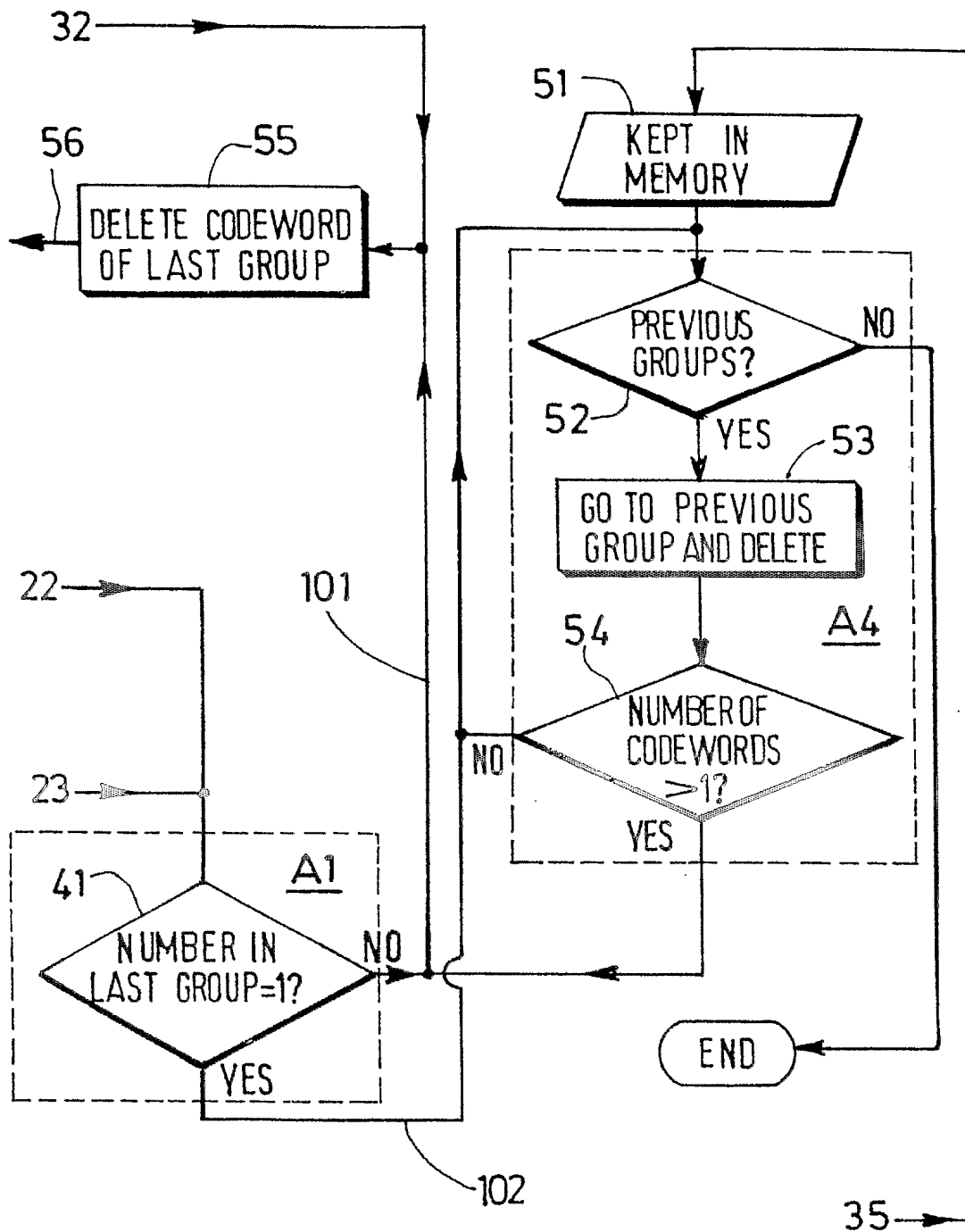


FIG. 11

d free	AL	TIME	CODE STRUCTURE
3	6,3494	42s	(1@ 4,-;1@ 5,-;5@ 6,3;6@ 7,3;4@ 8,3;4@ 9,3;1@ 10,-;1@ 12,-;2@ 13,-;1@ 14,-;2,1)
5	8,5061	3min	(1@ 6,-;2@ 7,5;3@ 8,5;4@ 9,5;4@ 10,5;5@ 11,5;5@ 12,5;2@ 13,5;3,2)
7	10,79	13min	(1@ 8,-;2@ 9,7;2@ 10,7;3@ 11,7;4@ 12,7;4@ 13,7;6@ 14,7;4@ 15,7;4,3)

FIG.12

d free	AL	TIME	CODE STRUCTURE
3	6,3530	17s	(1@ 4,-;1@ 5,-;5@ 6,3;6@ 7,3;4@ 8,3;3@ 9,3;2@ 10,4;2@ 11,3;1@ 12,-;1@ 13,-;2,1)
5	8,5061	45s	(1@ 6,-;2@ 7,5;3@ 8,5;4@ 9,5;4@ 10,5;5@ 11,5;5@ 12,5;2@ 13,5;3,2)
7	10,79	3min	(1@ 8,-;2@ 9,7;2@ 10,7;3@ 11,7;4@ 12,7;4@ 13,7;6@ 14,7;4@ 15,7;4,3)

FIG.13

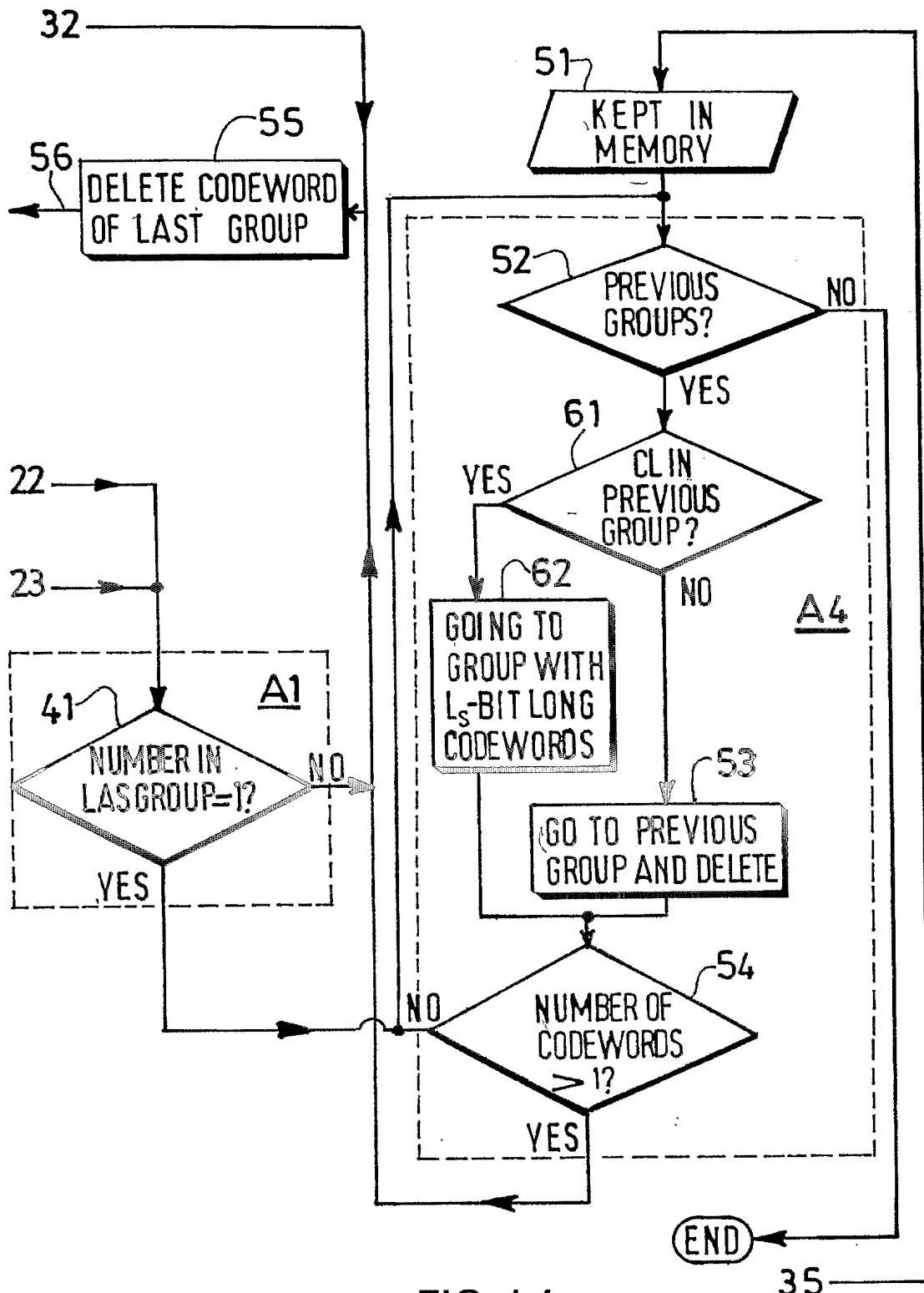


FIG.14



d tree	Ls->	4	5	6	7	8	9	10	11	12	13	14	15
3 (L <sub>1</sub> =4, L <sub>max</sub> =13)	AL Time	n° code	n° code	6,6356 0,017s	6,3650 0,08s	6,3550 0,5s	6,3530 3s	6,3530 6,7s	6,3530 11,5s	6,3530 17s	6,3530 17s	x x	x x
5 (L <sub>1</sub> =6, L <sub>max</sub> =13)	AL Time	x x	x x	n° code	8,5061 0,01s	8,5061 0,03s	8,5061 0,13s	8,5061 0,76s	8,5061 6,8s	8,5061 45s	8,5061 45s	x x	x x
7 (L <sub>1</sub> =8, L <sub>max</sub> =15)	AL Time	x x	x x	x x	x x	10,79 0,02s	10,79 0,07s	10,79 0,17s	10,79 0,6s	10,79 2s	10,79 12,8s	10,79 175s	10,79 175s

FIG.15

L	Ls->	5	6	7	8	9	10	11
9	AL Time	x x	x x	x x	x x	9,3728 1,7s	9,3728 100s	9,3728 3h
8	AL Time	x x	x x	x x	8,8570 0,7s	8,7575 21s	8,7537 18min	8,7537 32,5h
7	AL Time	x x	x x	n° code	8,5643 5s	8,3413 3min	8,3055 2h	8,2937 134h
6	AL Time	x x	n° code	n° code	8,4905 13s	8,1186 8min	8,0737 4h	not simulated
5	AL Time	n° code	n° code	n° code	8,3133 8s	8,0346 8min	7,9586 5h	7,9392 1Week

FIG.16

# INTERNATIONAL SEARCH REPORT

In ional Application No  
PCT/IB2004/000682

A. CLASSIFICATION OF SUBJECT MATTER  
IPC 7 H03M13/03 H03M13/35

According to International Patent Classification (IPC) or to both national classification and IPC

## B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)

IPC 7 H03M

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, WPI Data, INSPEC

## C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P, A	WO 03/047112 A (CHABBOUH SLIM ; LAMY CATHERINE (NL); KONINKL PHILIPS ELECTRONICS NV (N) 5 June 2003 (2003-06-05) the whole document	1-3
P, A	WO 03/043202 A (LAMY CATHERINE ; MEILHAC LISA (NL); KONINKL PHILIPS ELECTRONICS NV (NL) 22 May 2003 (2003-05-22) the whole document & US 5 745 504 A (BAANG KARL GOERAN) 28 April 1998 (1998-04-28)	1-3
A	WO 02/095955 A (LAMY CATHERINE ; POTHIER OLIVIER (NL); KONINKL PHILIPS ELECTRONICS NV) 28 November 2002 (2002-11-28) the whole document	1-3
	----- -/--	

☒ Further documents are listed in the continuation of box C.

☒ Patent family members are listed in annex.

\* Special categories of cited documents :

- \*A\* document defining the general state of the art which is not considered to be of particular relevance
- \*E\* earlier document but published on or after the international filing date
- \*L\* document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)
- \*O\* document referring to an oral disclosure, use, exhibition or other means
- \*P\* document published prior to the international filing date but later than the priority date claimed

- \*T\* later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention
- \*X\* document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone
- \*Y\* document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.
- \*Z\* document member of the same patent family

Date of the actual completion of the international search

1 July 2004

Date of mailing of the international search report

16/07/2004

Name and mailing address of the ISA

European Patent Office, P.B. 5818 Patentlaan 2  
NL - 2280 HV Rijswijk  
Tel. (+31-70) 340-2040, Tx. 31 651 epo nl,  
Fax: (+31-70) 340-3016

Authorized officer

Rydyger, K

## INTERNATIONAL SEARCH REPORT

International Application No  
PCT/IB2004/000682

## C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5 883 589 A (WADA MASAHIRO ET AL) 16 March 1999 (1999-03-16) the whole document -----	1-3
A	BUTTIGIEG V ET AL: "Variable-length error-correcting codes" IEE PROCEEDINGS : COMMUNICATIONS, INSTITUTION OF ELECTRICAL ENGINEERS, GB, vol. 147, no. 4, 15 August 2000 (2000-08-15), pages 211-215, XP006013980 ISSN: 1350-2425 the whole document page 214, column 1 - column 2 -----	1-3
A	CHABBOUH S ET AL: "A STRUCTURE FOR FAST SYNCHRONIZING VARIABLE-LENGTH CODES" IEEE COMMUNICATIONS LETTERS, IEEE SERVICE CENTER, PISCATAWAY, US, US, vol. 6, no. 11, November 2002 (2002-11), pages 500-502, XP001097285 ISSN: 1089-7798 the whole document figure 2a -----	1-3
A	VICTOR BUTTIGIEG: "Variable-length error-correcting codes" 1995, UNIVERSITY OF MANCHESTER , MANCHESTER , XP002286817 cited in the application page 152 - page 156 figure 6.6 -----	1-3

# INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No

PCT/JP2004/000682

Patent document cited in search report		Publication date	Patent family member(s)	Publication date
WO 03047112	A	05-06-2003	WO 03047112 A2	05-06-2003
WO 03043202	A	22-05-2003	WO 03043202 A1	22-05-2003
US 5745504	A	28-04-1998	AU 722461 B2	03-08-2000
			AU 3470997 A	14-01-1998
			CN 1228888 A ,B	15-09-1999
			EP 0908017 A1	14-04-1999
			JP 2000513164 T	03-10-2000
			WO 9750183 A1	31-12-1997
WO 02095955	A	28-11-2002	CN 1463500 T	24-12-2003
			EP 1397869 A1	17-03-2004
			WO 02095955 A1	28-11-2002
US 5883589	A	16-03-1999	JP 3267142 B2	18-03-2002
			JP 9232968 A	05-09-1997